

To cite this paper: Zhaohao Sun (2018) 10 Bigs: Big Data and Its Ten Big Characteristics,
BAIS No. 17010, PNG University of Technology

10 Bigs: Big Data and Its Ten Big Characteristics

Zhaohao Sun¹, Kenneth Strang², Rongping Li³

¹ Department of Business Studies

PNG University of Technology, Lae 411, PNG

zsun@dbb.unitech.ac.pg; zhaohao.sun@gmail.com

&

² School of Business & Economics

State University of New York, Plattsburgh at Queensbury, NY 12804, USA

kenneth.strang@plattsburgh.edu

&

School of Economics and Management

Hebei University of Science and Technology

Shijiazhuang, Hebei 050018, China

1965322@163.com

Abstract. This paper reveals ten big characteristics (10 Bigs) of big data and explores their non-linear interrelationships through presenting a unified framework of big data. The framework has three levels: fundamental level, technological level, and socio-economic level. The fundamental level has four big fundamental characteristics of big data. The technological level consists of three big technological characteristics of big data. The socioeconomic level has three big socioeconomic characteristics of big data. The paper looks at each level of the proposed framework from a service-oriented perspective. The proposed approach in this paper might facilitate the research and development of big data, big data analytics, data science, business intelligence, and business analytics.

Keywords: big data, big data analytics, business intelligence, data science, 10 Bigs.

1 Introduction

Big data has become one of the most important frontiers for innovation, research and development in computer science [1, 2], industry and business [3]. Big data has been also a key enabler of exploring business insights, business intelligence, and economics of services. This has drawn an unprecedented interest in industries, universities, governments and organizations [4]. Big data and its emerging technologies including big data analytics [5] have been not only making big changes in the way the business operate but also making traditional data analytics and business analytics bring about

new big opportunities for academia and enterprises [6]. However, some fundamentals of big data are still open for comprehending big data. For example,

- What are the characteristics of big data? and
- How can we understand the relationships among the characteristics of big data?

This paper will address these two issues. To address the first issue, this paper, different from the existing literature on big data, directly proposes and looks at 10 Bigs rather than 10 Vs [7] or 3 Vs [8] or 4 Vs [9] as the ten big characteristics of big data. This 10 Bigs consist of big volume, big velocity, big variety, big veracity, big intelligence, big analytics, big infrastructure, big service, big value, and big market. Then this paper explores the non-linear interrelationships among these 10 Bigs through presenting a unified framework. The framework reveals that 4 Bigs (i.e. big volume, big velocity, big variety, and big veracity) are fundamental characteristics of big data. Another 3 Bigs (i.e. big intelligence, big analytics, big infrastructure) are technological characteristics of big data. The remainder 3 Bigs (i.e. big service, big value, and big market) are socioeconomic characteristics of big data. The paper examines each level of the proposed framework from a service-oriented perspective.

The remainder of this paper is organized as follows. Section 2 overviews the characteristics of big data from an evolutionary Perspective. Section 3 proposes and looks at 10 Bigs as the ten big characteristics of big data. Section 4 examines the interrelationships among the 10 Bigs through presenting a unified framework. The final sections discuss the related work and end this paper with some concluding remarks and future work.

It should be noted that big might not be the best term to represent the characteristics of big data in some cases. However, it is only “big” of data that makes everyone draw much attention to big data, because everyone can engage in big data. Everyone can understand “big” well although not everyone can image how big the 100 Zettabytes (ZB) of data is.

2 Characteristics of Big Data

This section overviews the characteristics of big data from an evolutionary perspective.

McKinsey defines big data as “the datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze”. [10]. This definition makes McKinsey as one of the most important players usher big data in the next frontier for innovation, competition, and productivity. Gartner as another global consulting company defines big data as the “high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making” [4]. This definition, because of its marketing feature, has been most-frequently used in both analyst communities and academia communities for research and development of big data as a computing, science, and technology.

The characteristics of big data have been scattered in various publications. From an evolutionary perspective, Doug Laney of the META Group (now Gartner) uses 3 Ds: data volume, data velocity and data variety to represent the characteristics of data in e-commerce in 2001. Later these 3 Ds have been changed into 3 Vs (volume, velocity, and variety) which have been explained as three characteristics of big data [11, 8]. Furthermore, these 3 Vs have been extended first to 4 Vs as four characteristics of big data by adding veracity [12, 9], and then to 5 Vs as five characteristics of big data (volume, variety, velocity, veracity, value) [13, 14], finally to 10 Vs by adding another 5 Vs: validity, variability, venue, vocabulary and vagueness [7], although these 10 Vs are not the characteristics but 10 big challenges of big data in [7].

With the development of big data as a computing, big data computing, more Vs might be proposed to extend these 10 Vs, because case based reasoning (CBR) researchers have extended the R model of CBR from 4 Rs, 5 Rs [15] to 11 Rs [16]. However, these studies are a kind of linear thinking, because, from 3 Vs we can have 4 Vs, and then, mathematically, we can have infinite Vs (at least a big number) for understanding data rather than big data, and all the mentioned Vs are used to characterize data rather than big data. The importance of “big” has deeply been minded in a big data world. In this sense, the inception of big data does not belong to Laney [17], because he has not used a concept of big data at that time. Further, the researchers prefer to use esoteric concept rather than popular word to represent what they have done in their research, “big” is not a word in any jargon. The existing research on big data still prefers to Vs at the moment. Therefore, it is necessary to characterize “big” of the big data. If 3 Ds [17] can be changed into 3 Vs [8], then we can change 3 Vs to 3 Bigs for big volume, big variety, big velocity as the three “big” characteristics of big data to reflect the importance of “big” in big data. We can further extend 3 Bigs to 10 Bigs to examine ten big characteristics of big data and their interrelationships in terms of big data computing.

3 Ten Bigs as Big Characteristics of Big Data

This section proposes and examines ten big characteristics of big data with examples for each.

3.1 Big Volume

The big volume of big data reflects the size of the data set, which is typically in exabytes (EB) or ZB [18, 7] although many, like a child, might consider that 10 is big as an integer. Nowadays, many data-driven companies are working with petabytes (PB) of data daily. Google processes over 20 PB of data daily [19]. Walmart collects more than 2.5 PB of unstructured data from its 1 million customers every hour. Big data repositories for future generation parallel and distributed systems currently exceed EB and are increasing at a rapid speed in size [18]. The volume of big data has been increased from EB or ZB [1].

To overcome the challenge arising from big data with big volume, massively parallel computing platform is a practical choice, because its underlying principle is a

distribution of workload across many processors as well as storage and transportation of underlying data across a set of parallel storage units and streams [12, p. 32]. Hadoop has become the de facto standard for storing, processing and analyzing the data with big volume of hundreds of terabytes, and even PB [1]. For example, Hadoop clusters at Yahoo! span over 40,000 servers, and store 40 PB of application data [1].

3.2 Big Velocity

Big velocity is related to throughput and latency of data [12]. For throughput, big velocity means that at a big speed, data in and out from the networked systems in a real time [20, 3]. In other words, it is the high rate of data and information flowing into and out of interworked systems with real-time [7, 4].

Big velocity of big data is more important than big volume for many real-world applications [8]. In some developing countries, some internetworked systems cannot be running normally because of low speed, although big data are available there.

Latency is the other measure of velocity. Low latency is the requirement of modern businesses and individuals. For example, Turn (www.turn.com) is conducting its analytics in 10 milliseconds to place advertisements in online advertising platforms [12, p. 3]. IBM's InfoSphere Streams has been successfully used for low-latency, real-time analytics [12, p. 54].

3.3 Big Variety

Big variety means that big diversity or big different types of data sources with different structures from which it arrived, and the types of data available to everyone [20, 1, 21]. Big data can be classified into three types: structured, semi-structured, and unstructured at a higher level. The data stored in relational database systems like Oracle are structured. The data available on the web are unstructured. 80% of the world's data is unstructured [1, 12]. Blogs and tweets on social media are not structured data, because they contain a large amount of slang words, with a mix of languages in a multiethnic, multi-language environment [12, p. 41]. The big variety exists in the data on the Web. For example, in the WeChat world (<https://web.wechat.com/>), one can interact with his/her friends anywhere in the world using a variety of media such as text, sound clips, photos and videos.

The big variety brings about big challenge for storage, mining and analytics [1]. NoSQL databases have been used to overcome this challenge from big variety [1]. There are 225+ NoSQL databases on the market (<http://nosql-database.org/>, retrieved on 08 July 2016).

Big variety of big data is related to big venue and variability. Big venue of big data includes distributed big data, heterogeneous big data from multiple platforms, from different owners' systems, with different access and formatting requirements, private vs. public cloud [7]. Big variety of big data is also related to big variability which, e.g. consists of dynamic big data, evolving, spatiotemporal big data, time series big data, seasonal big data, and any other type of non-static data sources.

3.4 Big Veracity

Big veracity refers to the accuracy, truth and truthfulness of big data. Veracity was introduced by IBM researchers as the 4th V for characterizing big data [9]. The reality is that there exists big ambiguity, incompleteness, uncertainty of big data [9, 1, 7]. This might be the reason why vagueness has been considered as one of ten challenges of big data [7]. Accuracy and reliability are less controllable for many forms of big data, for example, Twitter posts a number of tweets with hash tags, abbreviations, typos and colloquial speech. Thus big veracity is a big characteristics of big data, this is particularly true in big data analytics for business decision-making [22]. Therefore, in order to get big veracity we have to use big data technology to remove the ambiguous, incomplete, uncertain data.

Fuzzy logic and fuzzy sets have developed significant methods and techniques to address ambiguity and incompleteness of data and therefore they will play an important role in overcoming big ambiguity and incompleteness of big data [23, 24].

3.5 Big Intelligence

Intelligence has been not only a lasting topic for computer science under the flagship of artificial intelligence (AI) and intelligence computing, but also an exciting topic for industries, organizations and businesses under the flagship of business intelligence (BI). AI has facilitated the development of intelligent services, intelligent manufacturing, and intelligent systems [25]. BI has promoted the improvement of competitiveness of business performance, and supported management decision making of organisations and produced the billionaire level enterprises such as Google and Facebook [26].

In the big data world, big intelligence particularly refers to big data intelligence. Big data intelligence is big data-driven intelligence, which can be defined as a set of ideas, technologies, systems and tools that can imitate the human intelligence related to big data management and processing, like business intelligence [22]. For example, intelligent methods for searching big data, visualizing knowledge mined from big data belongs to big data intelligence.

Big intelligence has been drawn increasing attention with the development of BI, big data analytics, and big data computing. Just as intelligence as a part of computing, big intelligence will be developed as an important part of big data computing (Wang, 2012). Global free WIFI [27] is also a big intelligence, because it meets the big expectation from billions of people [22].

3.6 Big Analytics

Big analytics is a brief representation of big data analytics or big data based analytics or big data-driven analytics [28]. Big analytics can be defined as the process of collecting, organizing and analyzing big data to discover and visualize patterns, knowledge, and intelligence as well as other information within the big data for supporting decision making [29]. The main components of big analytics include big descriptive analytics, big predictive analytics and big prescriptive analytics [28, 22].

- 1 Big descriptive analytics is descriptive analytics for big data [30, 24]. It addresses the problems such as what happened, and when, as well as what is happening. For example, web analytics for pay-per-click or email marketing data belongs to big data descriptive analytics [31].
- 2 Big data predicative analytics is predicative analytics for big data [29], which focuses on forecasting trends by addressing the problems such as what will happen, what's going to happen, what is likely to happen and why it will happen [30, 32, 8]. For example, big data predicative analytics can be used to predict where might be the next attack target of terrorists [22].
- 3 Big data prescriptive analytics is prescriptive analytics for big data [29], which addresses the problems such as what we should do, why we should do it and what should happen with the best outcome under uncertainty [28, p. 5]. For example, big data prescriptive analytics can be used to provide an optimal marketing strategy for an e-commerce company [22].

3.7 Big Infrastructure

Infrastructure refers to structures, systems, and facilities serving a country, city, or area, including the services and facilities necessary for its economy to function [33]. Big infrastructure, abbreviated from big data infrastructure refers to all the structures, technologies, systems, platforms, and facilities serving the big data processing in a country, city or area. The big data processing including big data ingestion, storing, analysis, modelling, data mining and knowledge discovery, visualization, simulation, and reporting [12] [22]. Currently, Apache Hadoop ecosystem has been considered as an important part of big infrastructure for processing big data [1]. Big infrastructure is a decisive factor for the utility of big data with applications, reflecting the level of national big data research and development, just as traffic infrastructure and network infrastructure as the criteria for national life standards and social development.

Although the price of the elements of computing machinery such as storage, memory, processing [8], bandwidth has been declining, big infrastructure of big data is still an important factor to differentiate developing countries from developed ones. This is also the reason why Chinese government launched big data infrastructure as a national development strategy in its 13th plan (2016-2020).

Global free WIFI is a big infrastructure in the big data world [27]. This might come into reality within ten years, just like free driving on the freeway in many countries.

3.8 Big Service

Services are playing a pivotal role in nations, organizations, businesses and individuals, service computing, social computing and cloud computing. Web services have been changing our work, life and thinking with the healthy development of service-centered society [34]. Big services are big data driven services or big data based on services. The big in big service means that it can provide big services to at least hundreds of millions people. For example, big data infrastructure services, cloud services, mobile services, big analytics services [22], social networking services are big services.

Big services, the web of services [34] and the Internet of services are interchangeably. The internet of services (IoS) is a big service covering all the services on the Internet based on big data. For example, online storage services are a kind of big services [35]. Big services are an emerging frontier for innovations, competences, and improving business performance of governments, organizations and enterprises.

3.9 Big Value

Big value indicates the importance and context of the big data [18]. It characterizes the big business value, ROI, and potential of big data to transform an organization to have more competitiveness in the global platform [7]. Big data has extremely big value for increasing productivity, efficiency and revenues, lowering costs and reducing risk in businesses and management [3, 21] [14]. For example, big data and big data analytics have brought big value for Facebook, Google, Amazon and made them become the top companies in the world [36]. It has also big value for scientific breakthrough in science and technology that have given and will give us a big number of opportunities to make great progresses in many fields. For example, many medical studies are based on big data to find better solutions to cure diabetes and hypertension [28].

Big value also implies that big data brings big social value. The big social value means that big data has been revolutionizing the society in terms of working, living and thinking [37]. The thinking methods have been changed because of big data, for example, learning heavily relies on books, library, schools, newspapers in the past whereas in the big data age, learning is based on search, learning as a search (LaaS), I learn using LaaS whenever and wherever I am.

3.10 Big Market

In the big data world, big market refers to big data driven market. Big market includes market of big data technology, systems, platforms, tools and services. According to IDC's research ((International Data Corporation), the big data technology and services market is expected to grow at a five year compound annual growth rate (CAGR) of more than 28% from US\$260.3 million in 2014 to US\$711.2 million in 2018 with the fastest growing segments being cloud Infrastructure, storage, discovery and analytics applications, and Networking Infrastructure [38].

The big market attracts the CEOs of big companies to make big decisions for developing big data and related technologies. For example, IDC predicts that the business analytics software market will grow at a 9.7% compound annual growth rate over the next five years from 2012 to 2017 [39].

4 Non-Linear Interrelationships among 10 Bigs of Big Data

This section looks at the non-linear interrelationships among these 10 big characteristics of big data.

The non-linear interrelationships can be represented using a unified framework as illustrated in Figure 1. This framework covers the 10 Bigs along with their interrelationships to one another nonlinearly. This framework consists of three levels: a bottom level, a middle level and a top level. They correspond to a fundamental level, technological level, and socio-economical level respectively.

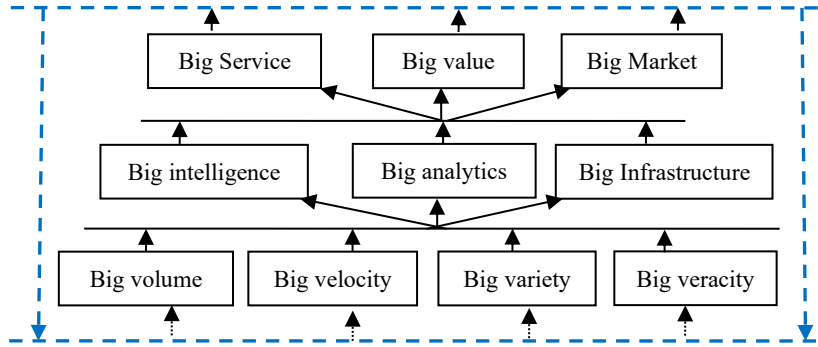


Figure 1. A unified framework of big data

In what follows, we will examine each level of the proposed framework from a service-oriented perspective [40].

At the bottom level, there are 4 Bigs: Big volume, big velocity, big variety, big veracity, they are four fundamental characteristics of big data from data science viewpoint in specific and computing viewpoint in general [22]. Integrating the definition of McKinsey on big data [10] and Gartner's definition of big data [4], this framework can define big data as a set of datasets, each of them satisfies that

1. It has big volume, big velocity, big variety and big veracity;
2. It is beyond the ability of typical ICT systems or tools to capture, store, manage, and analyse;
3. It demands cost-effective, innovative technologies of data processing for enhanced insight and decision making.

The main service providers on this level include computing scientists, and cloud services providers including Amazon, Google, and Facebook. The main service requestors on this level include everyone who can access to the Internet. The main service brokers on this level include big consulting groups like McKinsey, Gartner and IDC as well as public media [22].

Computing scientists are working on big intelligence, big analytics, and big infrastructure to meet the challenges of big data arising from the fundamental characteristics using ICT technologies, systems, and tools. This requires transforming from the bottom level to the middle level.

At the middle level, there are 3 Bigs: Big intelligence, big analytics, and big infrastructure. They are the three technological characteristics of big data. These 3 Bigs provide smart solutions and technologies to meet the challenges arising from the 4 Bigs

of the bottom level. The smart solutions and technologies include MapReduce NoSQL (Not only SQL) technology [5], MPP (Massively Parallel Processing) and In Memory database technologies [41, 42]. Apache Spark is a popular big data analytics platform for a number of enterprises [42, 43]. Spark provides dramatically increased large-scale data processing compared to Hadoop, and a NoSQL database for big data management [5, 43]. Apache Spark has provided Goldman Sachs with excellent big data analytics services [42].

The main service providers on this level include Google, Amazon, Tableau [42, 3] and other technology-driven companies, computing scientists and students. Big analytics service providers include developers, vendors, systems or software and other intermediaries that can provide big analytics services [29]. For example, Tableau as a software developer has been promoting big data and big analytics [42]. Amazon, Google and Microsoft, and Baidu are examples of big analytics service providers on the Web [29].

The main service requestors on this level include everyone who can access the Internet. Big analytics service requestors include organizations, governments and all level business decision makers such as CEO, CIO and CFO as well as managers. More generally, big analytics service requestors include people who like to acquire information based on analytical reports provided by big data analytics service providers [6]. Therefore, a person with smartphone receiving analytics services like GPS information is also a big data analytics service requestor [30, 29].

The main service brokers on this level are all the entities that facilitate the development of big intelligence, big analytics and big infrastructure, which include popular presses, traditional media and social media, consulting companies, scholars and university students, and so on [6]. Big consulting groups like McKinsey, Gartner and IDC as well as all the big data intermediaries [37] have played an important role as brokers at this level.

At the top level, there are 3 Bigs: Big service, big value, and big market. They are the three socioeconomic characteristics of big data. The socioeconomic consequence of dramatic development of big intelligence, big analytics and big infrastructure (at technological level) brings about big service, big value [10], and big market [44] to the society in a big data world. Further, big service, big value, and big market are dependent on each other closely. Big value from big data brings about big market [44], big market from big data brings about big services and big services bring about big value and big market, vice versa. All these statements have been proved and will be justified repeatedly in the big data world [8] [10] [28]. The main service providers on this level include big global industries such as manufacturing, healthcare, Finance and Insurance, social networking and all the companies related to Industry 4.0 [45], China Manufacturing 2025 [46] and soon on. The main service requestors on this level include everyone in the big data age. The main service brokers on this level include big consulting groups like McKinsey, Gartner and IDC and public media as well as the big data intermediaries [37].

The arrows with dash line represent that big service, big value and big market will bring about big impacts on the 4 big fundamental characteristics of big data at the bottom level. For example, it will increase the big volume of data from petabyte level to zettabyte level. It will improve the big velocity of big data in the Internet. It leads to

bigger variety and veracity of big data. On the other side, the arrows with dash line and other arrows form a cycle from and to big volume, big velocity, big variety and big veracity via big service, big value and big market to expedite the research and development of big data.

5 Related Work and Discussion

We have mentioned a number of scholarly researches on big data and its characteristics. In what follows, we will focus on related work and discussion on characteristics of big data.

Data volume, variety and velocity (3 Vs) were not introduced as the three characteristics of big data in 2001 but as a part of data management for e-commerce [17]. This implies that new characteristics of big data should be introduced in order to reflect the fast development of data since then. This has motivated us to delve the characteristics of big data.

Oweis et al [47] discuss 3 Vs, 4 Vs (3 Vs and veracity) and 5 Vs (4 Vs and value) as five characteristics of big data. Borne [7] extends the 5 Vs to 10 Vs of big data: volume, variety, velocity, veracity, value, validity, variability, venue, vocabulary and vagueness. However, Borne [7] has considered the 10 Vs as ten challenges of big data. Both of them have not examined the interrelationships among these mentioned Vs. This paper has proposed the 10 Bigs as ten big characteristics of big data and discussed their non-linear interrelationships from a service-oriented perspective. The 10 Bigs have not yet been wholly addressed in the literature although some of them have been studied separately. To our knowledge, big infrastructure, big service, big market, and big intelligence have not been formally introduced as the characteristics of big data.

Further, different from the 10 Vs of Borne [7], vagueness has been merged into veracity, because vagueness is related to uncertainty, ambiguity, incompleteness and fuzziness [23]. The big veracity of big data aims to eliminate this vagueness using big data technology [12]. Similarly, validity related to data quality, governance of data [7] can also be merged into veracity. Venue and variability of [7] have been merged in big variety. Further, this paper has not used some of the 10 Vs mentioned by [7], even if some Vs can be considered as characteristics of big data, but they are not big characteristics of big data, for example, vocabulary mentioned in [7], because vocabulary describing the data's structure, syntax, content, and provenance is a classic concept in database systems [5].

For the purpose of this research, we only focus on the 10 Bigs to provide a deep understanding of big data taking into account the limitation of space. If one likes to put more Bigs to characterize big data, then big opportunity [48], big challenge [42, 37, 3] and big impact are good choices for addition, because all these three have been closely associated with the discussion of big data [10] [8], although we have not put them in our 10 Bigs. For example, big data has brought about big challenges for our thinking, work, industry and society [28] and technologies [3]. Big data has also brought about big challenges to strategic decision making and organizational intelligence [49]. Big data has brought about new challenges to data quality management [12], because we have less control over the management of external data.

6 Conclusion

This paper revealed ten big characteristics of big data and explored their non-linear interrelationships through presenting a unified framework. The framework covers the 10 Bigs of big data and their non-linear interrelationships based on three levels: fundamental level, technological level, socio-economic level. This paper examined each level of the proposed framework from a service-oriented perspective.

10 Bigs of big data have been ushering in big revolution in existing sciences, technologies, methods and techniques in many fields. In the future work, we will delve the service-oriented framework of big data as a unified foundation for big data through analyzing the 10 Bigs of big data and viability of the proposed framework. We will also explore the applications of the framework in cloud services, and the Internet of things as well as the Internet of services through providing real-world cases in a big data world.

References

- [1]. B. Kumar, "An encyclopedic overview of 'big data' analytics," *International Journal of Applied Engineering Research*, vol. 10, no. 3, pp. 5681-5705, 2015.
- [2]. McKinsey, "The digital tipping point: McKinsey Global Survey results," June 2014. [Online]. Available: http://www.mckinsey.com/insights/business_technology/the_digital_tipping_point_mckinsey_global_survey_results. [Accessed 17 2014].
- [3]. C. P. Chen and C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data," *Information Sciences*, vol. 275, p. 314–347, 2014.
- [4]. Gartner, "Big data," 2016. [Online]. Available: <http://www.gartner.com/it-glossary/big-data/>.
- [5]. C. Coronel and S. Morris, *Database Systems: Design, Implementation, and Management* (11th edition), Boston: Cengage Learning, 2015.
- [6]. Z. Sun, K. Strang and J. Yearwood, "Analytics service oriented architecture for enterprise information systems," in *Proceedings of iiWAS2014, CONFENIS 2014, 4 - 6 Dec 14*, Hanoi, 2014.
- [7]. K. Borne, "Top 10 Big Data Challenges – A Serious Look at 10 Big Data V's," April 2014. [Online]. Available: <https://www.mapr.com/blog/top-10-big-data-challenges-%E2%80%93-serious-look-10-big-data-v%E2%80%99s>.
- [8]. A. McAfee and E. Brynjolfsson, "Big data: The management revolution," *Harvard Business Review*, pp. 61-68, 2012.

- [9]. IBM, "The Four V's of Big Data," 2015. [Online]. Available: <http://www.ibmdatahub.com/infographic/four-vs-bigdata>.
- [10]. McKinsey, "Big data: The next frontier for innovation, competition, and productivity," May 2011. [Online]. Available: <http://www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation>.
- [11]. C. Tsai, C. Lai, H. Chao and A. Vasilakos, "Big data analytics: a survey," *Journal of Big Data*, vol. 2, pp. 31-62, 2015.
- [12]. A. Sathi, Big data analytics: Disruptive technologies for changing the game, Boise, ID, USA: MC Press: IBM Corporation, 2013.
- [13]. DataCom, "DataCom 2015 International Conference on Big Data Intelligence and Computing," 3 11 2015. [Online]. Available: <http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=46304©ownerid=22598>.
- [14]. F.-Y. Wang, "A big-data perspective on AI: Newton, Merton, and Analytics Intelligence," *IEEE Intelligent Systems*, Sept/Oct, pp. 2-4, 2012 .
- [15]. Z. Sun and G. Finnie, Intelligent Techniques in E-Commerce: A Case-based Reasoning Perspective, Heidelberg Berlin: Springer-Verlag, 2004; 2010.
- [16]. D. Bridge, "The Virtue of Reward: Performance, Reinforcement and Discovery: Invited talk," 2005. [Online].
- [17]. D. Laney, "3D data management: controlling data volume, velocity, and variety, META Group, Tech. Rep.," 2001. [Online]. Available: <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>. [Accessed 27 10 2015].
- [18]. N. Vajjhala, K. Strang and Z. Sun, "Statistical modeling and visualizing of open big data using a terrorism case study," in *The International Conference on Open and Big Data (OBD 2015)*, 24-26 Aug., Rome, 2015.
- [19]. H. E. Pence, "What is big data and why is it important?," *Journal of Educational Technology Systems*, vol. 43, no. 2, pp. 159-171, 2015.
- [20]. J. Betser and D. Belanger, "Architecting the enterprise with big data analytics," in *Big Data and Business Analytics*, Boca Raton, FL, CRC Press , 2013, pp. 1-20.
- [21]. D. Loshin, Big Data Analytics: From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL and Graph, Amsterdam: Elsevier , 2013.
- [22]. Z. Sun, L. Sun and K. Strang, "Big data analytics services for enhancing business Intelligence," *Journal of Computer Information Systems*, p. DOI:10.1080/08874417.2016.1220239, 2016.
- [23]. L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8 , no. 3, p. 338–353, 1965.

- [24]. M. Kantardzic, *Data Mining: Concepts, Models, Methods, and Algorithms*, Hoboken, NJ: Wiley & IEEE Press, 2011.
- [25]. S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach* (3rd edition), Prentice Hall, 2010.
- [26]. R. Sabherwal and I. Becerra-Fernandez, *Business Intelligence: Practices, Technologies, and Management*, Wiley & Sons, 2011.
- [27]. M. Payton, "http://metro.co.uk/2015/06/25/google-wants-to-bring-free-wifi-to-the-world-and-its-starting-now-5265352/", 25 6 2015. [Online]. Available: <http://metro.co.uk/2015/06/25/google-wants-to-bring-free-wifi-to-the-world-and-its-starting-now-5265352/>. [Accessed 12 8 2016].
- [28]. M. Minelli, M. Chambers and A. Dhiraj, *Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses*, Wiley & Sons, 2013.
- [29]. Z. Sun, H. Zou and K. Strang, "Big Data Analytics as a Service for Business Intelligence," in *I3E2015, LNCS 9373*, Berlin, 2015.
- [30]. D. Delena and H. Demirkanb, "Data, information and analytics as services," *Decision Support Systems*, vol. 55, no. 1, p. 359–363, 2013.
- [31]. C. Cramer, "How Descriptive Analytics Are Changing Marketing," 19 May 2014. [Online]. Available: <http://www.miprofs.com/wp/descriptive-analytics-changing-marketing/>. [Accessed 6 7 2015].
- [32]. E. Turban and L. Volonino, *Information Technology for Management: Improving Strategic and Operational Performance* (8th Edition), Danvers, MA: Wiley & Sons, 2011.
- [33]. Wikipedia, "Infrastructure," May 2011. [Online]. Available: <https://en.wikipedia.org/wiki/Infrastructure>. [Accessed 20 7 2016].
- [34]. Z. Sun and J. Yearwood, "A theoretical foundation of demand-driven web services," in *Demand-Driven Web Services: Theory, Technologies, and Applications*, IGI-Global, 20014, pp. 1-25.
- [35]. C. C. Huang, P. Smith and Z. Sun, "Securing network for cloud services," in *Demand Driven Wen Services*, IGI-Global, 2014.
- [36]. D. Court, "Getting big impact from big data," January 2015. [Online]. Available: http://www.mckinsey.com/insights/business_technology/getting_big_impact_from_big_data. [Accessed 23 6 2015].
- [37]. V. Mayer-Schoenberger and K. Cukier, *Big Data: A Revolution that Will Transform How We Live, Work, and Think*, Houghton Mifflin Harcourt Publishing Company, 2013.
- [38]. S. Parker and A. T. Hira, "Australia Reaches a Tipping Point for Big Data Adoption, Says IDC Australia," 23 6 2015. [Online]. Available: <http://www.idc.com/getdoc.jsp?containerId=prAU25707715>.

- [39]. D. Vesset, B. McDonough, D. Schubmehl and M. Wardley, "Worldwide Business Analytics Software 2013–2017 Forecast and 2012 Vendor Shares (Doc # 241689)," 6 2013. [Online]. Available: <http://www.idc.com/getdoc.jsp?containerId=241689>. [Accessed 28 6 2014].
- [40]. M. P. Papazoglou and D. Georgakopoulos, "Service-oriented computing," *The Communications of the ACM*, vol. 46, no. 10, pp. 25-28, 2003.
- [41]. K. Bakshi, "Technologies for Big Data," in *Big Data Management, Technologies, and Applications*, IGI-Global, 2014, pp. 1-22.
- [42]. Tableau, "Top 8 Trends for 2016: Big Data," 2015. [Online]. Available: www.tableau.com/Big-Data.
- [43]. C. K. Reddy, "A survey of platforms for big data analytics," *Journal of Big Data (Springer)*, vol. 1, no. 8, pp. 1-20, 2014.
- [44]. IDC, "Big Data & Analytics," 2015. [Online]. Available: <https://www.idc.com/prodserv/4Pillars/bigdata>. [Accessed 1 8 2015].
- [45]. GTAI , "Industrie 4.0 Smart Manufacturing for the Future," 2014. [Online]. Available: http://www.gtai.de/GTAI/Content/EN/Invest/_SharedDocs/Downloads/GTAI/Brochures/Industries/industrie4.0-smart-manufacturing-for-the-future-en.pdf. [Accessed 13 8 2016].
- [46]. State Council, China, "Made in China 2025," 4 2016. [Online]. Available: <http://baike.baidu.com/item/%E4%B8%AD%E5%9B%BD%E5%88%B6%E9%80%A02025>. [Accessed 13 8 2016].
- [47]. N. E. Oweis and et al, "A Survey on Big Data, Mining: (Tools, Techniques, Applications and Notable Uses)," in *Intelligent Data Analysis and Applications*, Berlin, Springer International Publishing, 2015.
- [48]. U. Jovanović, A. Štimec and D. Vladušić, "Big data analytics: A critical review and some future directions," *International Journal of Business Intelligence and Data Mining*, vol. 10, no. 4, pp. 337-355, 2015.
- [49]. V. Morabito, *Big Data and Analytics: Strategic and Organizational Impacts*, Springer , 2015.